



PATENT ABSTRACTS OF JAPAN

(11) Publication number: **09073299 A**(43) Date of publication of application: **18.03.97**

(51) Int. Cl.

G10L 3/02
G10L 7/04
G11B 20/00
G11B 20/10
H04N 5/92
H04N 7/24

(21) Application number: **08169454**(22) Date of filing: **28.06.96**(30) Priority: **30.06.95 JP 07166385**(71) Applicant: **SANYO ELECTRIC CO LTD**

(72) Inventor:
YAMAUCHI HIDEKI
OKADA SHIGEYUKI
IIDA MASAYUKI
TANAKA KOJI

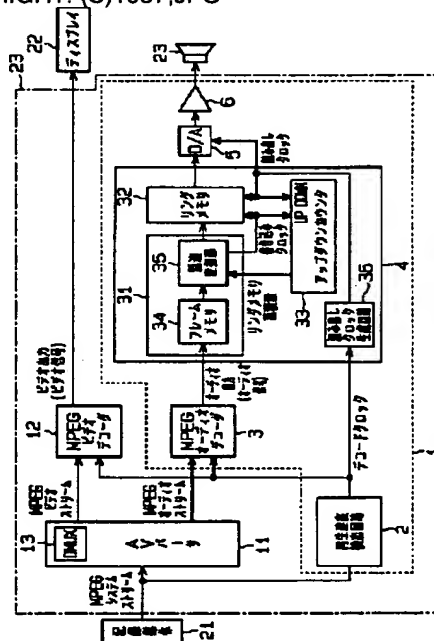
(54) **MPEG AUDIO REPRODUCING DEVICE AND
 MPEG REPRODUCING DEVICE**

COPYRIGHT: (C)1997,JPO

(57) Abstract:

PROBLEM TO BE SOLVED: To provide an MPEG audio reproducing device which reproduces audio signals that are easily understood during a variable speed reproducing.

SOLUTION: An MPEG audio reproducing device 1 consists of a reproducing speed detecting circuit 2, an MPEG audio decoder 3, a speech speed conversion processing circuit 4, a D/A converter 5 and an audio amplifier 6. Moreover, an MPEG reproducing device is provided with an audio-video purser (an AV purser) and an MPEG video decoder 12 in addition to the device 1. The circuit 4 consists of a DSP 31, a ring memory 32 and an up-down counter 33. The circuit 4 expands the time length of the voice segment inputted during a high speed reproducing and reduces the time length of each silence interval. During a low speed reproducing, the time length of each voice segment is expanded, the time length of each silence interval is reduced or each silence interval is deleted, each voice segment is connected together and inserted into a silence interval.



(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平 9 - 7 3 2 9 9

(43) 公開日 平成9年(1997)3月18日

(51) Int. Cl. ^a	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L	3/02		G 1 0 L	3/02 A
	7/04			7/04 G
G 1 1 B	20/00		G 1 1 B	20/00 Z
	20/10	3 2 1		20/10 3 2 1 Z
H 0 4 N	5/92	7736-5 D	H 0 4 N	5/92 H
			O L	
審査請求 未請求 請求項の数 1 1			(全 2 7 頁) 最終頁に続く	

(21) 出願番号 特願平8-169454

(22) 出願日 平成8年(1996)6月28日

(31) 優先権主張番号 特願平7-166385

(32) 優先日 平7(1995)6月30日

(33) 優先権主張国 日本 (J P)

(71) 出願人 000001889

三洋電機株式会社

大阪府守口市京阪本通2丁目5番5号

(72) 発明者 山内 英樹

大阪府守口市京阪本通2丁目5番5号 三洋電機株式会社内

(72) 発明者 岡田 茂之

大阪府守口市京阪本通2丁目5番5号 三洋電機株式会社内

(72) 発明者 飯田 正幸

大阪府守口市京阪本通2丁目5番5号 三洋電機株式会社内

(74) 代理人 弁理士 恩田 博宣

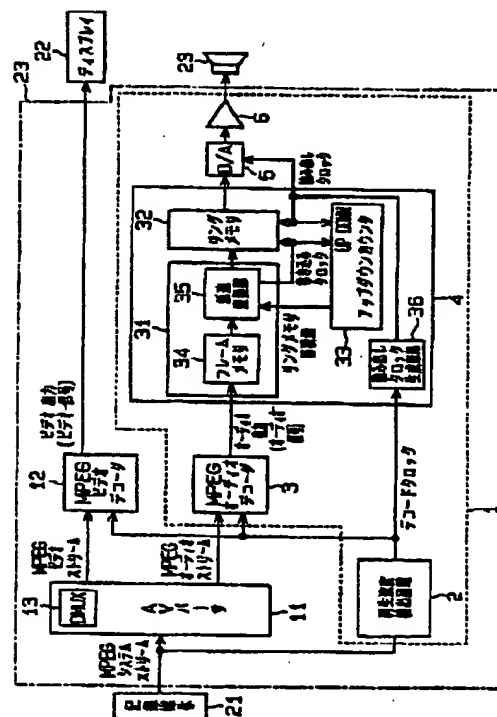
最終頁に続く

(54) 【発明の名称】 MPEGオーディオ再生装置およびMPEG再生装置

(57) 【要約】

【課題】 可変速再生時においても聞き易い音声を再生することが可能なMPEGオーディオ再生装置を提供する。

【解決手段】 MPEGオーディオ再生装置 1 は再生速度検出回路 2、MPEGオーディオデコーダ 3、話速変換処理回路 4、D/Aコンバータ 5、オーディオアンプ 6 から成る。また、MPEG再生装置はMPEGオーディオ再生装置 1 に加え、オーディオビデオパーサ (AVパーサ)、MPEGビデオデコーダ 12 を備える。話速変換処理回路 4 はDSP 31、リングメモリ 32、アップダウンカウンタ 33 から構成される。話速変換処理回路 4 は、高速再生時には入力される音声区間の時間長さを伸長すると共に各無音区間の時間長さを短くするようにし、低速再生時には、各音声区間の時間長さを伸長すると共に各無音区間の時間長さを短くするか、または、各無音区間を削除して各音声区間をつなぎ合わせた後に無音区間を挿入する。



1

【特許請求の範囲】

【請求項1】 MPEGオーディオデコーダ(3)と、その出力に対して話速変換処理を行う話速変換処理手段(2, 4)とを備えたMPEGオーディオ再生装置。

【請求項2】 MPEGオーディオデコーダ(3)と、その出力に対して話速変換処理を行う話速変換処理手段(2, 4)と、MPEGビデオデコーダ(3)とを備えたMPEG再生装置。

【請求項3】 記録媒体(21)から読み出されたMPEGオーディオストリームをMPEGオーディオパートに準拠してデコードし、オーディオ信号を生成するMPEGオーディオデコーダ(3)と、オーディオ信号に対して話速変換処理を行う話速変換処理手段(2, 4)とを備え、

話速変換処理手段は、オーディオストリームのビットレートが通常時よりも大きい場合には、再生される音声のピッチを通常の再生時とほぼ同一にし、且つ、再生される話速を通常の再生時に近づけるように話速変換処理を行い、オーディオストリームのビットレートが通常時よりも小さい場合には、音声区間の途切れが目立たなくなるように話速変換処理を行うMPEGオーディオ再生装置。

【請求項4】 記録媒体(21)から読み出されたMPEGオーディオストリームをMPEGオーディオパートに準拠してデコードし、オーディオ信号を生成するMPEGオーディオデコーダ(3)と、オーディオ信号に対して話速変換処理を行う話速変換処理手段(2, 4)とを備え、

話速変換処理手段は、オーディオストリームのビットレートが通常時よりも大きい場合には、再生される音声区間の時間長さを伸長すると共に各無音区間の時間長さを短くするようにして話速変換処理を行い、オーディオストリームのビットレートが通常時よりも小さい場合には、再生される各音声区間の時間長さを伸長すると共に各無音区間の時間長さを短くするか、または、各無音区間を削除して各音声区間をつなぎ合わせた後に無音区間を挿入するようにして話速変換処理を行うMPEGオーディオ再生装置。

【請求項5】 請求項3または請求項4に記載のMPEGオーディオ再生装置において、話速変換処理手段(2, 4)は、オーディオ信号を蓄積するリングメモリ(32)と、リングメモリの蓄積量を検出する検出手段(33)とを備え、リングメモリの蓄積量に応じて音声区間の時間長さの圧縮伸長率を調整するMPEGオーディオ再生装置。

【請求項6】 請求項5に記載のMPEGオーディオ再生装置において、話速変換処理手段(2, 4)は、オーディオ信号の音声区間と無音区間とを判別する音声

2

判別部(41)と、

無音区間の削除処理または挿入処理を行う無音削除挿入部(42)と、

リングメモリ(32)の蓄積量に基づいて音声区間の圧縮伸長処理を行うことで圧縮伸長率を調整する時間軸圧縮伸長部(43)とを備えたMPEGオーディオ再生装置。

【請求項7】 請求項3～6のいずれか1項に記載のMPEGオーディオ再生装置(1)と、

10 記録媒体(21)から読み出されたMPEGビデオストリームをMPEGビデオパートに準拠してデコードし、ビデオ信号を生成するMPEGビデオデコーダ(12)とを備えたMPEG再生装置。

【請求項8】 請求項5または請求項6に記載のMPEGオーディオ再生装置(1)と、

記録媒体(21)から読み出されたMPEGビデオストリームをMPEGビデオパートに準拠してデコードし、ビデオ信号を生成するMPEGビデオデコーダ(12)と、

20 リングメモリ(32)に書き込まれる以前のオーディオ信号に、時刻に関する情報としてのインデックス信号を付加するインデックス付加回路(51)と、

リングメモリ(32)から読み出されたオーディオ信号に付加されているインデックス信号を検出し、そのインデックス信号から得られる時刻情報と現在の時刻情報とから、話速変換処理手段(2, 4)における信号遅延時間を検出し、その検出された遅延時間を示す信号をMPEGビデオデコーダ(12)へ供給するインデックス検出回路(52)とを備え、

30 MPEGビデオデコーダ(12)は、前記遅延時間を示す信号に基づいて自己の動作のタイミングを制御するMPEG再生装置。

【請求項9】 請求項6に記載のMPEGオーディオ再生装置(1)と、

記録媒体(21)から読み出されたMPEGビデオストリームをMPEGビデオパートに準拠してデコードし、ビデオ信号を生成するMPEGビデオデコーダ(12)と、

40 音声判別部(41)の処理結果と、オーディオストリームのビットレートとに基づいて、話速変換処理手段

(2, 4)における信号遅延時間を検出し、その検出された遅延時間を示す信号をMPEGビデオデコーダ(12)へ供給する遅延時間検出回路(53)とを備え、MPEGビデオデコーダ(12)は、前記遅延時間を示す信号に基づいて自己の動作のタイミングを制御するMPEG再生装置。

【請求項10】 請求項6に記載のMPEGオーディオ再生装置(1)と、

50 記録媒体(21)から読み出されたMPEGビデオストリームをMPEGビデオパートに準拠してデコードし、

ビデオ信号を生成するMPEGビデオデコーダ(12)と、
リングメモリ(32)の蓄積量に基づいて、話速変換処理済みのオーディオ信号とビデオ信号との同期を得るための制御信号を生成し、その制御信号をMPEGビデオデコーダ(12)へ供給する制御回路(54)とを備え、
MPEGビデオデコーダ(12)は、前記制御信号に基づいて自己の動作のタイミングを制御するMPEG再生装置。

【請求項11】 請求項6に記載のMPEGオーディオ再生装置(1)と、
記録媒体(21)から読み出されたMPEGビデオストリームをMPEGビデオパートに準拠してデコードし、ビデオ信号を生成するMPEGビデオデコーダ(12)と、
音声判別部(41)および時間軸圧縮伸長部(43)の処理結果に基づいて、話速変換処理手段(2, 4)における信号遅延時間を検出し、その検出された遅延時間を示す信号をMPEGビデオデコーダ(12)へ供給する遅延時間検出回路(55)とを備え、
MPEGビデオデコーダ(12)は、前記遅延時間を示す信号に基づいて自己の動作のタイミングを制御するMPEG再生装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明はMPEG(Moving Picture Expert Group)オーディオ再生装置およびMPEG再生装置に係り、詳しくは、話速変換機能を備えたMPEGオーディオ再生装置およびMPEG再生装置に関するものである。

【0002】

【従来の技術】マルチメディアで扱われる情報は、膨大な量で且つ多種多様であり、これらの情報を高速に処理することがマルチメディアの実用化を図る上で必要となってくる。情報を高速に処理するためには、データの圧縮・伸長技術が不可欠となる。そのようなデータの圧縮・伸長技術として「MPEG」方式が挙げられる。このMPEG方式は、ISO(International Organization for Standardization)/IEC(International Electrotechnical Commission)傘下のMPEG委員会(ISO/IEC JTC1/SC29/WG11)によって標準化されつつある。

【0003】MPEGは3つのパートから構成されている。パート1の「MPEGシステムパート」(ISO/IEC IS 11172 Part1:Systems)では、ビデオデータとオーディオデータの多重化構造(マルチプレクス・ストラクチャ)および同期方式が規定される。パート2の「MPEGビデオパート」(ISO/IEC IS 11172 Part2:Video)では、ビデオデータの高性能符号化方式およびビデオデータのフォーマットが規定される。パート3の「MPEG

オーディオパート」(ISO/IEC IS 11172 Part3:Audio)では、オーディオデータの高性能符号化方式およびオーディオデータのフォーマットが規定される。

【0004】MPEGビデオパートで取り扱われるビデオデータは動画に関するものであり、その動画は1秒間に数十個(例えば、30個)のフレーム(静止画、コマ)によって構成されている。ビデオデータは、シーケンス(Sequence)、GOP(Group Of Pictures)、ピクチャ、スライス(Slice)、マクロブロック(Macroblock)、ブロックの順に6層の階層構造から成る。

【0005】また、MPEGには主にエンコードレートの違いにより、現在のところ、MPEG-1、MPEG-2の2つの方式がある。MPEG-1においてフレームはピクチャに対応している。MPEG-2においては、フレームまたはフィールドをピクチャに対応させることもできる。フィールドは、2個で1つのフレームを構成している。ピクチャにフレームが対応している構造はフレーム構造と呼ばれ、ピクチャにフィールドが対応している構造はフィールド構造と呼ばれる。

【0006】MPEGでは、フレーム間予測と呼ばれる圧縮技術を用いる。フレーム間予測は、フレーム間のデータを時間的な相関に基づいて圧縮する。フレーム間予測では双方向予測が行われる。双方向予測とは、過去の再生画像(または、ピクチャ)から現在の再生画像を予測する順方向予測と、未来の再生画像から現在の再生画像を予測する逆方向予測とを併用することである。

【0007】この双方向予測は、Iピクチャ(Intra-Picture)、Pピクチャ(Predictive-Picture)、Bピクチャ(Bidirectionally predictive-Picture)と呼ばれる3つのタイプのピクチャを規定している。Iピクチャは、過去や未来の再生画像とは無関係に、独立して生成される。Pピクチャは順方向予測(過去のIピクチャまたはPピクチャからの予測)により生成される。Bピクチャは双方向予測により生成される。双方向予測においてBピクチャは、以下に示す3つの予測のうちいずれか1つにより生成される。①順方向予測;過去のIピクチャまたはPピクチャからの予測、②逆方向予測;未来のIピクチャまたはPピクチャからの予測、③双方向予測;過去および未来のIピクチャまたはPピクチャからの予測。そして、これらI、P、Bピクチャがそれぞれエンコードされる。つまり、Iピクチャは過去や未来のピクチャが無くとも生成される。これに対し、Pピクチャは過去のピクチャが無いと生成されず、Bピクチャは過去または未来のピクチャが無いと生成されない。

【0008】フレーム間予測では、まず、Iピクチャが周期的に生成される。次に、Iピクチャよりも数フレーム先のフレームがPピクチャとして生成される。このPピクチャは、過去から現在への一方向(順方向)の予測により生成される。続いて、Iピクチャの前、Pピクチャの後に位置するフレームがBピクチャとして生成され

る。このBピクチャを生成するとき、順方向予測、逆方向予測、双方向予測の3つの中から最適な予測方法が選択される。一般的に連続した動画では、現在の画像とその前後の画像とは良く似ており、異なっているのはその一部分に過ぎない。そこで、前のフレーム（例えば、Iピクチャ）と次のフレーム（例えば、Pピクチャ）とは同じであると仮定し、両フレーム間に変化があればその差分（Bピクチャ）のみを抽出して圧縮する。これにより、フレーム間のデータを時間的な相関に基づいて圧縮することができる。

【0009】MPEGビデオパートに準拠してエンコードされたビデオデータのデータ列（ビットストリーム）は、MPEGビデオストリーム（以下、ビデオストリームと略す）と呼ばれる。また、MPEGオーディオパートに準拠してエンコードされたオーディオデータのデータ列は、MPEGオーディオストリーム（以下、オーディオストリームと略す）と呼ばれる。そして、ビデオストリームとオーディオストリームは、MPEGシステムパートに準拠して時分割多重化され、1本のデータ列としてのMPEGシステムストリーム（以下、システムストリームと略す）となる。システムストリームはマルチプレックスストリームとも呼ばれる。

【0010】MPEGパートにおけるエンコードからデコードまでの流れは、以下のようになっている。MPEGシステムエンコーダ（以下、システムエンコーダと略す）は、ビデオデータとオーディオデータのそれぞれを連係を保ちながら別個にエンコードを行い、ビデオストリームとオーディオストリームを生成する。次に、MPEGシステムエンコーダに装備されたマルチプレクサ

（MUX；Multiplexer）は、伝送媒体または記録媒体のフォーマットに適合するように、ビデオストリームとオーディオストリームの多重化を行い、システムストリームを生成する。そのシステムストリームは、伝送媒体を介してMUXから伝送されるか、または記録媒体に記録される。

【0011】MPEGシステムデコーダ（以下、システムデコーダと略す）に装備されたデマルチプレクサ（DMUX；DeMultiplexer）は、システムストリームをビデオストリームとオーディオストリームに分離する。次に、システムデコーダは各ストリームを個別にデコードして、ビデオのデコード出力（以下、ビデオ出力という）とオーディオのデコード出力（以下、オーディオ出力という）を生成する。ビデオ出力はディスプレイへ出力され、ディスプレイで動画が再生される。オーディオ出力はD/A（Digital/Analog）コンバータおよびオーディオアンプを介してスピーカへ出力され、スピーカから音声再生される。

【0012】ところで、MPEG-1は主にビデオCD（Compact Disc）、CD-ROM（CD-Read Only Memory）、DVD（Digital Video Disc）などの記録媒体を

用いた蓄積メディアに対応しており、MPEG-2はMPEG-1をも含む幅広い範囲のアプリケーションに対応している。

【0013】蓄積メディアにおいては、以下に示す2つの可変速再生が要求される。①動画を通常（標準）の再生速度より高速で再生（以下、高速再生という）する機能。②動画を通常の再生速度より低速で再生（以下、低速再生という）する機能。高速再生機能は、例えば、ユーザが短時間に動画を見るために早送り再生を行う際や、見たい動画を探索するために早送り再生または早送り逆転再生を行う際に用いられる。また、低速再生機能は、例えば、ユーザが動画を注意深く見る際などに用いられる。

【0014】記録媒体から読み出されたシステムストリームのビットレートは、読み出し速度に対応したものになる。従って、高速再生を行うには記録媒体からシステムストリームを高速で読み出し、低速再生を行うには記録媒体からシステムストリームを低速で読み出す。例えば、記録媒体としてビデオCDやDVDを用いた場合には、ビデオCDやDVDの回転速度を通常の再生時（標準再生時）よりも速くしたり遅くしたりすることで、システムストリームを所望の速度で読み出すようにする。

【0015】

【発明が解決しようとする課題】従来、MPEGにおいては、前記したような動画の可変速再生については検討されていたものの、音声の可変速再生については何らの検討もなされていなかった。

【0016】オーディオストリームのビットレートはシステムストリームのそれと同一である。そのため、動画の高速再生時には、オーディオストリームのビットレートも大きくなり、再生される音声の音程（ピッチ）が上がるのに加えて、発声速度（話速）が速くなる。また、動画の低速再生時には、オーディオストリームのビットレートも小さくなり、再生される音声のピッチは変化しないものの、音声途切れ途切れになる。このように、動画の可変速再生時には、音声聞き苦しいものになるという問題があった。

【0017】ところで、近年、ピッチを変化させることなく話速を任意に制御する話速変換技術の開発が進められており、本出願人もVTRやテープレコーダに利用可能な話速変換処理LSIを既に開発している（特開平7-192992号公報（G11B20/02）、日経エレクトロニクス1994年11月21日号（No.622）P.93～98、参照）。しかし、話速変換技術をMPEGに利用する試みはなされていない。

【0018】また、音声と動画（映像）の同期生成においては、「リップシンク」を考慮する必要がある。リップシンクとは、ディスプレイに映し出される人物の口の動きと、スピーカから発声される音声との同期がとれていないことをいう。口の動きより音声の方が早くなった

り、逆に遅くなったりする状態をリップシンクにずれがあるという。リップシンクのずれが人間の聴覚の許容範囲を外れると、視聴者は違和感を覚える。一般に、音声動画より遅れることによって生じるリップシンクのずれとして許容できる時間は、約50～250msであるといわれている。

【0019】本発明は上記要求を満足するためになされたものであって、以下の目的を有するものである。

〔1〕可変速再生時においても自然で聞き易い音声を再生することが可能なMPEGオーディオ再生装置を提供する。

【0020】〔2〕上記〔1〕のMPEGオーディオ再生装置とMPEGビデオデコーダとを備えたMPEG再生装置を提供する。

〔3〕上記〔1〕のMPEGオーディオ再生装置とMPEGビデオデコーダとを備え、音声と動画との時間ずれを低減することが可能なMPEG再生装置を提供する。

【0021】

【課題を解決するための手段】請求項1に記載の発明は、MPEGオーディオデコーダ(3)と、その出力に対して話速変換処理を行う話速変換処理手段(2、4)とを備えたことをその要旨とする。

【0022】請求項2に記載の発明は、MPEGオーディオデコーダ(3)と、その出力に対して話速変換処理を行う話速変換処理手段(2、4)と、MPEGビデオデコーダ(3)とを備えたことをその要旨とする。

【0023】請求項3に記載の発明は、記録媒体(21)から読み出されたMPEGオーディオストリームをMPEGオーディオパートに準拠してデコードし、オーディオ信号を生成するMPEGオーディオデコーダ

(3)と、オーディオ信号に対して話速変換処理を行う話速変換処理手段(2、4)とを備え、話速変換処理手段は、オーディオストリームのビットレートが通常時よりも大きい場合には、再生される音声のピッチを通常の再生時とほぼ同一にし、且つ、再生される話速を通常の再生時に近づけるように話速変換処理を行い、オーディオストリームのビットレートが通常時よりも小さい場合には、音声区間の途切れが目立たなくなるように話速変換処理を行うことをその要旨とする。

【0024】請求項4に記載の発明は、記録媒体(21)から読み出されたMPEGオーディオストリームをMPEGオーディオパートに準拠してデコードし、オーディオ信号を生成するMPEGオーディオデコーダ

(3)と、オーディオ信号に対して話速変換処理を行う話速変換処理手段(2、4)とを備え、話速変換処理手段は、オーディオストリームのビットレートが通常時よりも大きい場合には、再生される音声区間の時間長さを伸長すると共に各無音区間の時間長さを短くするようにして話速変換処理を行い、オーディオストリームのビットレートが通常時よりも小さい場合には、再生される各

音声区間の時間長さを伸長すると共に各無音区間の時間長さを短くするか、または、各無音区間を削除して各音声区間をつなぎ合わせた後に無音区間を挿入するようにして話速変換処理を行うことをその要旨とする。

【0025】請求項5に記載の発明は、請求項3または請求項4に記載のMPEGオーディオ再生装置において、話速変換処理手段(2、4)は、オーディオ信号を蓄積するリングメモリ(32)と、リングメモリの蓄積量を検出する検出手段(33)とを備え、リングメモリの蓄積量に応じて音声区間の時間長さの圧縮伸長率を調整することをその要旨とする。

【0026】請求項6に記載の発明は、請求項5に記載のMPEGオーディオ再生装置において、話速変換処理手段(2、4)は、オーディオ信号の音声区間と無音区間とを判別する音声判別部(41)と、無音区間の削除処理または挿入処理を行う無音削除挿入部(42)と、リングメモリ(32)の蓄積量に基づいて音声区間の圧縮伸長処理を行うことで圧縮伸長率を調整する時間軸圧縮伸長部(43)とを備えたことをその要旨とする。

【0027】請求項7に記載の発明は、請求項3～6のいずれか1項に記載のMPEGオーディオ再生装置

(1)と、記録媒体(21)から読み出されたMPEGビデオストリームをMPEGビデオパートに準拠してデコードし、ビデオ信号を生成するMPEGビデオデコーダ(12)とを備えたことをその要旨とする。

【0028】請求項8に記載の発明は、請求項5または請求項6に記載のMPEGオーディオ再生装置(1)と、記録媒体(21)から読み出されたMPEGビデオストリームをMPEGビデオパートに準拠してデコードし、ビデオ信号を生成するMPEGビデオデコーダ(12)と、リングメモリ(32)に書き込まれる以前のオーディオ信号に、時刻に関する情報としてのインデックス信号を付加するインデックス付加回路(51)と、リングメモリ(32)から読み出されたオーディオ信号に付加されているインデックス信号を検出し、そのインデックス信号から得られる時刻情報と現在の時刻情報とから、話速変換処理手段(2、4)における信号遅延時間を検出し、その検出された遅延時間を示す信号をMPEGビデオデコーダ(12)へ供給するインデックス検出回路(52)とを備え、MPEGビデオデコーダ(12)は、前記遅延時間を示す信号に基づいて自己の動作のタイミングを制御することをその要旨とする。

【0029】請求項9に記載の発明は、請求項6に記載のMPEGオーディオ再生装置(1)と、記録媒体(21)から読み出されたMPEGビデオストリームをMPEGビデオパートに準拠してデコードし、ビデオ信号を生成するMPEGビデオデコーダ(12)と、音声判別部(41)の処理結果と、オーディオストリームのビットレートとに基づいて、話速変換処理手段(2、4)における信号遅延時間を検出し、その検出された遅延時間

を示す信号をMPEGビデオデコーダ(12)へ供給する遅延時間検出回路(53)とを備え、MPEGビデオデコーダ(12)は、前記遅延時間を示す信号に基づいて自己の動作のタイミングを制御することをその要旨とする。

【0030】請求項10に記載の発明は、請求項6に記載のMPEGオーディオ再生装置(1)と、記録媒体(21)から読み出されたMPEGビデオストリームをMPEGビデオパートに準拠してデコードし、ビデオ信号を生成するMPEGビデオデコーダ(12)と、リングメモリ(32)の蓄積量に基づいて、話速変換処理済みのオーディオ信号とビデオ信号との同期を得るための制御信号を生成し、その制御信号をMPEGビデオデコーダ(12)へ供給する制御回路(54)とを備え、MPEGビデオデコーダ(12)は、前記制御信号に基づいて自己の動作のタイミングを制御することをその要旨とする。

【0031】請求項11に記載の発明は、請求項6に記載のMPEGオーディオ再生装置(1)と、記録媒体(21)から読み出されたMPEGビデオストリームをMPEGビデオパートに準拠してデコードし、ビデオ信号を生成するMPEGビデオデコーダ(12)と、音声判別部(41)および時間軸圧縮伸長部(43)の処理結果に基づいて、話速変換処理手段(2, 4)における信号遅延時間を検出し、その検出された遅延時間を示す信号をMPEGビデオデコーダ(12)へ供給する遅延時間検出回路(55)とを備え、MPEGビデオデコーダ(12)は、前記遅延時間を示す信号に基づいて自己の動作のタイミングを制御することをその要旨とする。

【0032】

【発明の実施の形態】

(第1実施形態) 以下、本発明を具体化した第1実施形態を図面に従って説明する。

【0033】図1に、本実施形態のブロック回路図を示す。本実施形態のMPEGオーディオ再生装置1は、再生速度検出回路2、MPEGオーディオデコーダ3、話速変換処理回路4、D/Aコンバータ5、オーディオアンプ6から構成されている。尚、各回路2~6は1チップのLSIに搭載することもできる。

【0034】また、本実施形態のMPEG再生装置23は、MPEGオーディオ再生装置1に加え、オーディオビデオパーサ(AVパーサ)11、MPEGビデオデコーダ12を備えている。

【0035】話速変換処理回路4は、例えば、DSP(Digital Signal Processor)31、リングメモリ32、アップダウンカウンタ33、読み出しクロック生成回路36を備えている。尚、話速変換処理回路4の動作については、前記文献(日経エレクトロニクス 1994年11月21日号(No. 622) P. 93~98.)に詳述されている。

【0036】再生速度検出回路2は、ビデオCDやDV

Dなどの記録媒体21から読み出されたMPEGシステムストリームのビットレートに対応したデコードクロックを生成する。そのデコードクロックは各回路12, 3, 4へ出力される。

【0037】AVパーサ11は、デマルチプレクサ(DMUX)13を備えており、記録媒体21から読み出されたMPEGシステムストリームを入力する。DMUX13は、システムストリームをMPEGビデオストリームとMPEGオーディオストリームに分離する。ビデオストリームはビデオデコーダ12へ出力され、オーディオストリームはオーディオデコーダ3へ出力される。

【0038】ビデオデコーダ12は、MPEGビデオパートに準拠してビデオストリームをデコードし、ビデオ出力(以下、ビデオ信号という)を生成する。そのビデオ信号はディスプレイ22へ出力され、ディスプレイ22で動画が再生される。

【0039】オーディオデコーダ3は、MPEGオーディオパートに準拠してオーディオストリームをデコードし、ディジタル信号のオーディオ出力(以下、オーディオ信号という)を生成する。そのオーディオ信号は話速変換処理回路4へ出力される。話速変換処理回路4において信号処理されたオーディオ信号はD/Aコンバータ5によってD/A変換された後、オーディオアンプ6で増幅されてスピーカ23へ送られる。そして、スピーカ23から音声再生される。

【0040】記録媒体21から読み出されたシステムストリームのビットレートは、読み出し速度に対応したものになる。また、各回路3, 4, 12の動作はデコードクロックによって規定される。

【0041】従って、ビデオデコーダ12は、システムストリームのビットレートに対応したビデオ信号を生成する。すなわち、システムストリームのビットレートが、通常の再生時(標準再生時)よりも大きければディスプレイ22では動画が高速再生され、通常の再生時よりも小さければディスプレイ22では動画が低速再生される。

【0042】また、オーディオデコーダ3は、システムストリームのビットレートに対応したオーディオ信号を生成する。すなわち、システムストリームのビットレートが、通常の再生時よりも大きければオーディオ信号のビットレートも大きくなり、通常の再生時よりも小さければオーディオ信号のビットレートも小さくなる。

【0043】ところで、ビデオ信号とオーディオ信号とは、通常の再生時において同期生成されるようになっていく。DSP31は、フレームメモリ34および話速変換部35から構成されている。フレームメモリ34は、適宜なフレーム数分(例えば、2フレーム分)のオーディオ信号を記憶する。話速変換部35は、フレームメモリ34に記憶されたオーディオ信号に対してフレーム単位で話速変換処理を行い、話速変換処理済みのオーディ

オ信号（以下、データという）を生成する。尚、1フレームは、適宜な数（例えば、200個）のサンプリングデータから構成される。

【0044】フレームメモリ34の内部は、2つの領域（以下、A領域、B領域と記載して区別する）に分けられている。オーディオデコーダ3から出力されたオーディオ信号がB領域に書き込まれるのと同時に、A領域に蓄積されている1フレーム分のオーディオ信号が読み出されて話速変換部35へ転送される。そして、B領域に1フレーム分のオーディオ信号が蓄積されると、今度は、B領域に蓄積された1フレーム分のオーディオ信号が読み出されて話速変換部35へ転送され、それと同時に、オーディオデコーダ3から出力されたオーディオ信号がA領域に書き込まれる。

【0045】話速変換部35の生成したデータは、話速変換部35が生成した書き込みクロックに従ってリングメモリ32に書き込まれる。リングメモリ32は、例えば、FIFO（First-In-First-Out）構成のRAM（Random Access Memory）から成る。

【0046】読み出しクロック生成回路36は、デコードクロックに従って読み出しクロックを生成する。リングメモリ32に蓄積されたデータは、読み出しクロックに従って読み出され、その読み出されたデータはD/Aコンバータ5へ出力される。D/Aコンバータ5は、読み出しクロックをサンプリング周波数として用いる。

【0047】書き込みクロックはアップダウンカウンタ33のアップカウント入力端子UPに入力され、読み出しクロックはアップダウンカウンタ33のダウンカウント入力端子DOWNに入力される。アップダウンカウンタ33は、書き込みクロックの総数と読み出しクロックの総数との差をカウントする。そのカウント値は、リングメモリ32の蓄積量に対応する。つまり、アップダウンカウンタ33は、書き込みクロックと読み出しクロックとに基づいて、リングメモリ32の蓄積量を検出する。そのリングメモリ32の蓄積量は話速変換部35へ出力される。

【0048】図2に、話速変換部35に内部構成を示す。話速変換部35は、音声判別部41、無音削除挿入部42、時間軸圧縮伸長部43から構成されている。

【0049】音声判別部41は、フレームメモリ34から読み出されたオーディオ信号が、音声区間（音声が存在している区間）か、または、無音区間（音声が存在していない区間）かを判別する。尚、人間が発声する音声以外の背景雑音は無音区間として取り扱う。

【0050】無音削除挿入部42は、音声判別部41の判別した無音区間に対して、その無音区間の削除処理、または、新たな無音区間の挿入処理を行う。時間軸圧縮伸長部43は、音声判別部41の判別した音声区間に対して、リングメモリ32の蓄積量に基づいて圧縮処理または伸長処理を行う。

【0051】また、各部42、43は、その処理内容に対応した書き込みクロックを生成する。次に、高速再生時における話速変換部35の動作について説明する。

【0052】オーディオデコーダ3から出力されるオーディオ信号のビットレートは、オーディオストリームのそれと同一になる。従って、高速再生時には、通常の再生時に比べて、オーディオ信号のビットレートが大きくなる。通常の再生時よりもビットレートの大きなオーディオ信号をそのままD/Aコンバータ5へ送った場合、通常の再生時に比べて、スピーカ23から再生される音声のピッチは上がり話速は速くなる。

【0053】そこで、話速変換部35において、スピーカ23から再生される音声のピッチを通常の再生時とほぼ同一にし、且つ、スピーカ23から再生される話速を通常の再生時に近づけるように話速変換処理を行う。

【0054】すなわち、無音削除挿入部42は、音声判別部41の判別した無音区間の継続長を算出し、その継続長が所定長以上の場合は無音区間を削除する。また、時間軸圧縮伸長部43は、音声判別部41の判別した音声区間に対して、例えば、自己相関法を用いてピッチ抽出を行い、抽出したピッチ波形に対して圧縮処理を行う。その結果、高速再生時において、オーディオ信号のビットレートが大きくなった場合に、スピーカ23から再生される音声区間の時間長さは伸長される。

【0055】尚、時間軸圧縮伸長部43における圧縮処理に際しては、無音区間の状態とリングメモリ32の蓄積量とに応じて動的に圧縮率を変化させる。例えば、同一のピッチ周期をもつ3周期波形を2周期波形に圧縮することで、2/3倍の圧縮（圧縮率：2/3）を得る。具体的には、3周期波形から、時間軸方向で前にある2周期波形と、後ろにある2周期波形とをそれぞれ切り出す。そして、前の2周期波形に単調減少する三角窓関数を、後ろの2周期波形に単調増加する三角窓関数をそれぞれ乗じる。この二つの波形を加算することで出力波形を得る。

【0056】また、0.9倍の圧縮（圧縮率：0.9）を得るには、例えば、10周期波形から9周期波形に圧縮する。この場合は、先頭の3周期波形に対して同様の処理を施す。つまり、入力10周期波形のうち、先頭の3周期波形を除いた7周期波形は処理に使わない。

【0057】このM周期波形からN周期波形に圧縮する組み合わせを色々と用意しておくことで、多種類の圧縮率を得る。ところで、無音区間が短い場合、圧縮率が低い（圧縮の度合いが大きい）とリングメモリ32がオーバーフローする恐れがある。これを防ぐためには、リングメモリ32の蓄積量に応じて、時間軸圧縮伸長部43における圧縮率を動的に変化させればよい。また、背景雑音が存在する場合、音声区間やピッチの抽出誤りが生じる。これを防ぐためには、音声判別部41における音声区間の検出レベルを雑音信号に応じて変化させればよ

い。

【0058】次に、低速再生時における話速変換部35の動作について、図3および図4に従って説明する。図3に、通常の再生時および0.5倍速再生時において再生される音声の例を示す。

【0059】低速再生時には、通常の再生時に比べて、オーディオ信号のビットレートが小さくなる。そのため、方法1に示すように、通常の再生時よりもビットレートの小さなオーディオ信号をそのままD/Aコンバータ5へ送った場合、通常の再生時に比べて、スピーカ23から再生される音声のピッチは変化しないものの、音声の途切れ途切れになる。つまり、各音声区間（「あ」「い」「う」「え」）の時間長さは通常の再生時のそれと変わらず、全く音の存在していない無音区間が各音声区間の間に挿入されるため、音声の途切れ途切れになり、ユーザは聴感上違和感を覚える。

【0060】そこで、話速変換部35において、方法2または方法3に示すように話速変換処理を行う。尚、MPEGオーディオでは、低速再生時に音声のピッチが変化しないため、高速再生時のように時間軸圧縮伸長部43においてピッチを変える処理を行う必要はない。

【0061】（方法2）方法2では、時間軸圧縮伸長部43において各音声区間の長さを伸長させ、それと共に、無音削除挿入部42において各無音区間の長さを短くすることで、音声の途切れを目立たなくする。

【0062】尚、時間軸圧縮伸長部43において音声区間の長さを伸長させるには、音声判別部41の判別した音声区間に対して、例えば、自己相関法を用いてピッチ抽出を行い、抽出したピッチ波形に対して伸長処理を行う。例えば、同一のピッチ周期をもつ2周期波形を3周期波形に伸長することで、 $3/2$ 倍の伸長（伸長率： $3/2$ ）を得る。また、同一のピッチ周期をもつ3周期波形を4周期波形に伸長することで、 $4/3$ 倍の伸長（伸長率： $4/3$ ）を得る。その結果、低速再生時において、オーディオ信号のビットレートが小さくなった場合に、スピーカ23から再生される音声区間の時間長さは伸長される。

【0063】このとき、音声区間を伸長し過ぎると、音声区間が間延びして聞こえるため、音声の途切れは目立たなくなるものの、やはり不自然になる。これを防止するには、通常の再生時における音声区間の長さ $L1$ に対して、低速再生時における音声区間の長さ $L2$ を、例えば、以下の式に示すように設定する。

【0064】 $L2/L1 \leq 1.4$

尚、上記式は0.5倍速再生時だけでなく、あらゆる倍率の低速再生時に適用できる。ここで、時間軸圧縮伸長部43における音声区間の伸長率は一定値にしてもよく、以下の①②に示すように可変にしてもよい。

【0065】①リングメモリ32の蓄積量に対応して音声区間の伸長率を動的に変化させる。無音区間が短い場

合、音声区間の伸長率が大い（伸長の度合いが大い）とリングメモリ32がオーバーフローする恐れがある。これを防ぐためには、音声区間の伸長率を小さくすればよい。

【0066】②音声のピッチ変化に対応して音声区間の伸長率を動的に変化させる。つまり、図4に示すように、音声のピッチ変化に対応して音声区間の伸長率を変化させることで、話速を変化させる。この場合、音声の聞き易さをさらに向上させることができる。尚、音声のピッチ変化に対応して音声区間の伸長率を変化させることで話速を変化させる技術は公知である（信学技報 SP92-56, HC92-33(1992-09), P. 49~56. 参照）。

【0067】（方法3）方法3では、無音削除挿入部42において、各無音区間を削除して各音声区間をつなぎ合わせた後で、音声区間に続いて新たに無音区間を挿入することで、音声の途切れを目立たなくする。尚、挿入する無音区間は、以下の①~③のいずれであってもよい。

【0068】①全く音の存在しない無音区間。

②視聴者が違和感を覚えないような白色雑音を含む無音区間。尚、そのような白色雑音は、予め作成して別メモリ（図示略）に記憶しておく。

【0069】③音声判別部41において無音区間と判別したオーディオ信号を別メモリ（図示略）に保持しておき、それを無音区間として挿入する。

このように、本実施形態によれば、以下の作用および効果を得ることができる。

【0070】（1）話速変換処理回路4を設けることで、高速再生時において、スピーカ23から再生される音声のピッチを通常の再生時とほぼ同一にし、且つ、スピーカ23から再生される話速を通常の再生時に近づけることが可能になり、自然で聞き易い音声を再生することができる。

【0071】ところで、 m 倍速再生時（ $m > 1$ ）には、オーディオストリームおよびデコードクロックのビットレートは通常の再生時の m 倍になる。このとき、話速変換部35から出力されるデータのビットレートを通常の再生時とほぼ同一になるようにすれば、再生される音声のピッチを通常の再生時とほぼ同一にすることができる。すなわち、話速変換部35においてビットレートを $m \rightarrow 1$ に変換すれば、再生される音声のピッチは通常の再生時とほぼ同一になる。

【0072】（2）話速変換処理回路4を設けることで、低速再生時において再生される音声の途切れを目立たなくすることが可能になり、自然で聞き易い音声を再生することができる。

【0073】ところで、上記方法2と方法3とを、以下の(1)(2)に示すように併用してもよい。

(1) MPEGオーディオ再生装置1のユーザが、方法2と方法3とを任意に切り換え選択できるようにする。こ

のようにすれば、個々のユーザの聴覚特性に合わせる事が可能になり、ユーザにとって聞き易い音声を再生することができる。

(2) 低速再生の倍率に対応して方法 2 と方法 3 とが自動的に切り換え選択されるようにする。例えば、1~0.5 倍速再生時には方法 3 が選択され、0.5 倍速以下の再生時には方法 2 が選択されるようにする。このようにすれば、再生速度に応じて、自然な音声を再生することができる。

【0074】(3) 各回路 2~6 を 1 チップの LSI に搭載した場合には、MPEG オーディオ再生装置 1 を小型化することができる。

(第 2 実施形態) 以下、本発明を具体化した第 2 実施形態を図面に従って説明する。尚、本実施形態において、第 1 実施形態と同じ構成部品については符号を等しくしてその詳細な説明を省略する。

【0075】図 5 に、本実施形態の要部ブロック回路図を示す。本実施形態において、第 1 実施形態と異なるのは、インデックス付加回路 5 1 およびインデックス検出回路 5 2 が設けられている点だけである。

【0076】インデックス付加回路 5 1 は、フレームメモリ 3 4 の前段（すなわち、MPEG オーディオデコーダ 3 と話速変換処理回路 4 の間）に設けられている。インデックス付加回路 5 1 は、デコードクロックに従って、オーディオデコーダ 3 の生成したオーディオ信号に一定周期でインデックス信号を付加する。そのインデックス信号が付加されたオーディオ信号は、フレームメモリ 3 4 へ出力される。

【0077】インデックス検出回路 5 2 は、リングメモリ 3 2 から読み出されたデータに付加されているインデックス信号を検出し、そのインデックス信号から得られる時刻情報と現在時刻とから、話速変換処理回路 4 が信号処理に要する時間 Δt を算出し、その時間 Δt に関する検出信号をビデオデコーダ 1 2 へ供給する。ビデオデコーダ 1 2 は、その時間 Δt に関する検出信号に従って、自己の動作のタイミングを制御する。

【0078】このように、本実施形態によれば、第 1 実施形態の作用および効果に加えて、以下の作用および効果を得ることができる。

(1) 前記したように、ビデオデコーダ 1 2 の生成するビデオ信号と、オーディオデコーダ 3 の生成するオーディオ信号とは、通常の再生時において同期生成されるようになっている。そのため、オーディオデコーダ 3 と D/A コンバータ 5 の間に話速変換処理回路 4 を設けると、話速変換処理回路 4 における信号処理に要する時間分（すなわち、話速変換処理回路 4 における遅延時間分）だけ、オーディオ信号が遅延することになる。

【0079】そこで、インデックス付加回路 5 1 を用いて、フレームメモリ 3 4 へ入力されるオーディオ信号に予め一定周期でインデックス信号を付加する。インデッ

クス検出回路 5 2 は、リングメモリ 3 2 から読み出されたデータに付加されているインデックス信号を検出し、話速変換処理回路 4 が信号処理に要する時間 Δt を算出し、その時間 Δt に関する検出信号をビデオデコーダ 1 2 へ供給する。ビデオデコーダ 1 2 は、その時間 Δt に関する検出信号に従って、自己の動作のタイミングを制御する。また、インデックス検出回路 5 2 が次にインデックス信号を検出したとき、ビデオデコーダ 1 2 は、そのときに算出された時間と前回算出された時間との差だけ、自己の動作のタイミングを遅らせたり早めたりする。

【0080】その結果、話速変換処理回路 4 における遅延時間に関係なく、リングメモリ 3 2 から読み出されたデータ（すなわち、話速変換処理済みのオーディオ信号）とビデオ信号との同期をとることができる。

【0081】(2) 上記 (1) より、スピーカ 2 3 で再生される音声と、ディスプレイ 2 2 で再生される動画との時間ずれを低減することが可能になり、リップシンクのずれを人間の聴覚の許容範囲内にすることができる。

【0082】(3) オーディオ信号に付加されたインデックス信号は、無音削除挿入部 4 2 によって削除されることがある。しかし、インデックス信号を付加する周期を短くして、オーディオ信号に十分な数のインデックス信号を付加しておけば、そのインデックス信号の内のいくつかが無音削除挿入部 4 2 によって削除されたとしても、リングメモリ 3 2 から読み出されたデータには一定数以上のインデックス信号が残ることになる。その残ったインデックス信号により、上記 (1) の作用および効果を得ることができる。

【0083】(第 3 実施形態) 以下、本発明を具体化した第 3 実施形態を図面に従って説明する。尚、本実施形態において、第 2 実施形態と同じ構成部品については符号を等しくしてその詳細な説明を省略する。

【0084】図 6 に、本実施形態の要部ブロック回路図を示す。本実施形態において、第 2 実施形態と異なるのは、インデックス付加回路 5 1 が、フレームメモリ 3 4 と音声判別部 4 1 の間に設けられている点だけである。インデックス付加回路 5 1 は、デコードクロックに従って、フレームメモリ 3 4 から読み出されたオーディオ信号に一定周期でインデックス信号を付加する。そのインデックス信号が付加されたオーディオ信号は、音声判別部 4 1 へ出力される。

【0085】前記したように、フレームメモリ 3 4 が 2 フレーム分のオーディオ信号を蓄積する場合、フレームメモリ 3 4 の記憶容量は、例えば、0.8 K バイト程度あれば十分である。このように、フレームメモリ 3 4 の記憶容量が小さい場合には、話速変換処理回路 4 における遅延時間に比べて、フレームメモリ 3 4 における書き込み動作および読み出し動作に要する時間（すなわち、フレームメモリ 3 4 における遅延時間）は僅かであり、

無視しても差し支えない。

【0086】従って、本実施形態によれば、第2実施形態と同様の作用および効果を得ることができる。

(第4実施形態)以下、本発明を具体化した第4実施形態を図面に従って説明する。尚、本実施形態において、第2実施形態と同じ構成部材については符号を等しくしてその詳細な説明を省略する。

【0087】図7に、本実施形態の要部ブロック回路図を示す。本実施形態において、第2実施形態と異なるのは、インデックス付加回路51が、音声判別部41と無音削除挿入部42および時間軸圧縮伸長部43との間にそれぞれ設けられている点だけである。インデックス付加回路51は、デコードクロックに従って、音声判別部41における信号処理が済んだオーディオ信号に一定周期でインデックス信号を付加する。そのインデックス信号が付加されたオーディオ信号は、無音削除挿入部42および時間軸圧縮伸長部43へ出力される。

【0088】前記したように、フレームメモリ34の記憶容量が小さい場合には、話速変換処理回路4における遅延時間に比べて、フレームメモリ34における遅延時間は僅かであり、無視しても差し支えない。

【0089】また、音声判別部41における信号処理に要する時間（すなわち、音声判別部41における遅延時間）は、話速変換処理回路4における遅延時間に比べて僅かであり、無視しても差し支えない。

【0090】従って、本実施形態によれば、第2実施形態と同様の作用および効果を得ることができる。

(第5実施形態)以下、本発明を具体化した第5実施形態を図面に従って説明する。尚、本実施形態において、第2実施形態と同じ構成部材については符号を等しくしてその詳細な説明を省略する。

【0091】図8に、本実施形態の要部ブロック回路図を示す。本実施形態において、第2実施形態と異なるのは、インデックス付加回路51が、無音削除挿入部42および時間軸圧縮伸長部43とリングメモリ32との間に設けられている点だけである。インデックス付加回路51は、デコードクロックに従って、各部42、43における信号処理が済んだオーディオ信号に一定周期でインデックス信号を付加する。そのインデックス信号が付加されたオーディオ信号は、リングメモリ32へ出力される。

【0092】前記したように、フレームメモリ34の記憶容量が小さい場合には、話速変換処理回路4における遅延時間に比べて、フレームメモリ34における遅延時間は僅かであり、無視しても差し支えない。

【0093】また、各部41～43における信号処理に要する時間（すなわち、各部41～43における遅延時間）は、話速変換処理回路4における遅延時間に比べて僅かであり、無視しても差し支えない。

【0094】つまり、話速変換処理回路4における遅延

時間は、主に、リングメモリ32における書き込み動作および読み出し動作に要する時間（すなわち、リングメモリ32における遅延時間）によって決定される。

【0095】従って、本実施形態によれば、第2実施形態と同様の作用および効果を得ることができる。また、本実施形態によれば、第2実施形態のようにオーディオ信号に付加されたインデックス信号が無音削除挿入部42によって削除されることがない。そのため、付加したインデックス信号が全て活用され、インデックス信号の数を減らすことが可能になることから、インデックス付加回路51の回路規模を小さくすることができる。

【0096】(第6実施形態)以下、本発明を具体化した第6実施形態を図面に従って説明する。尚、本実施形態において、第1実施形態と同じ構成部材については符号を等しくしてその詳細な説明を省略する。

【0097】図9に、本実施形態の要部ブロック回路図を示す。本実施形態において、第1実施形態と異なるのは、遅延時間検出回路53が設けられている点だけである。前記したように、音声判別部41は、フレームメモリ34から読み出されたオーディオ信号が、音声区間か又は無音区間かを判別する。つまり、音声判別部41の処理結果には、オーディオ信号に音声が含まれているか否かという情報が含まれている。

【0098】また、デコードクロックは、システムストリームのビットレートに対応している。つまり、デコードクロックには、予めオーディオ信号の圧縮伸長率の情報が含まれている。

【0099】そこで、遅延時間検出回路53は、オーディオ信号に音声が含まれているか否かという情報と圧縮伸長率の情報とに基づいて、話速変換処理回路4における遅延時間を検出し、その検出信号をビデオデコーダ12へ供給する。ビデオデコーダ12は、遅延時間検出回路53の検出信号に基づいて、自己の動作のタイミングを制御する。その結果、話速変換処理回路4における遅延時間に関係なく、リングメモリ32から読み出されたデータ（すなわち、話速変換処理済みのオーディオ信号）とビデオ信号との同期をとることができる。

【0100】このように、本実施形態によれば、第2実施形態と同様の効果を得ることができる。

(第7実施形態)以下、本発明を具体化した第7実施形態を図面に従って説明する。尚、本実施形態において、第1実施形態と同じ構成部材については符号を等しくしてその詳細な説明を省略する。

【0101】図10に、本実施形態の要部ブロック回路図を示す。本実施形態において、第1実施形態と異なるのは、制御回路54が設けられている点だけである。制御回路54は、アップダウンカウンタ33の検出したリングメモリ32の蓄積量に基づいて、ビデオデコーダ12の動作速度を制御するための制御信号を生成し、その制御信号をビデオデコーダ12へ供給する。ビデオデコ

ーダ12は、制御回路54の制御信号に基づいて、自己の動作のタイミングを制御する。その結果、リングメモリ32から読み出されたデータと、ビデオデコーダ12の生成するビデオ信号との同期をとることができる。

【0102】前記したように、話速変換処理回路4における遅延時間は、主にリングメモリ32における遅延時間によって決定される。リングメモリ32における遅延時間は、その蓄積量と相関関係があり、蓄積量が大きくなるほど遅延時間も大きくなる。従って、リングメモリ32の蓄積量に基づいてビデオデコーダ12の動作速度を制御すれば、リングメモリ32から読み出されたデータ（すなわち、話速変換処理済みのオーディオ信号）とビデオ信号との同期をとることができる。

【0103】このように、本実施形態によれば、第2実施形態と同様の効果を得ることができる。

（第8実施形態）以下、本発明を具体化した第8実施形態を図面に従って説明する。尚、本実施形態において、第1実施形態と同じ構成部材については符号を等しくしてその詳細な説明を省略する。

【0104】図11に、本実施形態の要部ブロック回路図を示す。本実施形態において、第1実施形態と異なるのは、遅延時間検出回路55が設けられている点だけである。

【0105】前記したように、音声判別部41の処理結果には、オーディオ信号に音声が含まれているか否かという情報が含まれている。また、時間軸圧縮伸長部43の処理結果には、オーディオ信号の圧縮伸長率の情報が含まれている。

【0106】そこで、遅延時間検出回路55は、オーディオ信号に音声が含まれているか否かという情報と圧縮伸長率の情報とに基づいて、話速変換処理回路4における遅延時間を検出し、その検出信号をビデオデコーダ12へ供給する。ビデオデコーダ12は、遅延時間検出回路55の検出信号に基づいて、自己の動作のタイミングを制御する。その結果、話速変換処理回路4における遅延時間に関係なく、リングメモリ32から読み出されたデータ（すなわち、話速変換処理済みのオーディオ信号）とビデオ信号との同期をとることができる。

【0107】このように、本実施形態によれば、第2実施形態と同様の効果を得ることができる。図12に、可変速再生機能を備えたMPEGビデオデコーダ12の要部ブロック回路を示す。

【0108】MPEGビデオデコーダ12は、ビットバッファ202、ピクチャヘッダ検出回路203、MPEGビデオデコードコア回路（以下、デコードコア回路と略す）204、可変閾値オーバーフロー判定回路（以下、判定回路と略す）205、ピクチャスキップ回路206、制御コア回路207から構成されている。尚、各回路203～207は1チップのLSIに搭載することもできる。

【0109】制御コア回路207は各回路2～6を制御する。AVパーサ11から転送されてきたMPEGビデオストリームはビットバッファ202へ入力される。

【0110】ビットバッファ202はFIFO構成のRAMから成るリングメモリによって構成され、転送されてくるビデオストリームをそのまま順次蓄積する。ピクチャヘッダ検出回路203は、ビットバッファ202に蓄積されたビデオストリームの各ピクチャの先頭に付くピクチャヘッダを検出し、その各ピクチャヘッダに規定されているピクチャのタイプ（I、P、B）を検出する。

【0111】制御コア回路207は、ピクチャヘッダ検出回路203の検出結果と後記する判定回路205の判定結果とに基づいて、ビットバッファ202から1フレーム期間毎に適宜なピクチャ分のビデオストリームを読み出す。尚、ビットバッファ202から読み出されたビデオストリームは、読み出された後もビットバッファ202にそのまま残される。

【0112】ビットバッファ202から読み出された各ピクチャは、ピクチャスキップ回路206を介してデコードコア回路204へ転送される。デコードコア回路204は、各ピクチャをMPEGビデオパートに準拠してデコードし、各ピクチャ毎のビデオ信号を生成する。

【0113】ピクチャスキップ回路206は、制御コア回路207の制御に従って各ノード206a、206b側への接続が切り換えられる。そして、ピクチャスキップ回路206がノード206a側に接続されると、ビットバッファ202から読み出されたピクチャはそのままデコードコア回路204へ転送される。また、ノード206b側に接続されると、ビットバッファ202から読み出されたピクチャはビットバッファ202へ転送されずにスキップされる。その結果、デコードコア回路204へ転送されるピクチャは、ピクチャスキップ回路206によってスキップされた分だけピクチャ単位で間引かれる。

【0114】判定回路205は、再生速度検出回路2の生成したデコードクロックに基づいてビットバッファ202の占有量 B_m の閾値 B_{thn} を設定し、ビットバッファ202の占有量 B_m と閾値 B_{thn} とを比較する。尚、判定回路205では、再生速度検出回路2の生成した実際のデコードクロックの周波数と、通常の再生時のデコードクロックの周波数との比を求め、その比を再生速度の倍率 n とする。従って、2倍速再生時には倍率 $n=2$ となり、閾値 $B_{thn}=B_{th2}$ となる。また、通常の再生時には倍率 $n=1$ となり、閾値 $B_{thn}=B_{th1}$ となる。

【0115】そして、判定回路205は、ビットバッファ202の占有量 B_m が閾値 B_{thn} を越えない場合には、ビットバッファ202がオーバーフローする恐れがなく正常であると判定する。この場合、制御コア回路207は、ビットバッファ202から1ピクチャ分のビデオ

21

オストリームを読み出す。そして、制御コア回路207は、ピクチャスキップ回路206をノード206a側に接続し、そのビットバッファ202から読み出されたピクチャをデコードコア回路204へ転送させる。

【0116】また、判定回路205は、ビットバッファ202の占有量 B_m が閾値 B_{thn} を越えた場合には、ビットバッファ202がオーバーフローする恐れがあると判定する。この場合、制御コア回路207は、ビットバッファ202の占有量 B_m が閾値 B_{thn} を下回るまで、ビットバッファ202から適宜なピクチャ分のビデオストリームを読み出す。そして、制御コア回路207は、ピクチャスキップ回路206をノード206b側に接続し、そのビットバッファ202から読み出された適宜なピクチャ分のビデオストリームを全てスキップさせる。

【0117】図13に、ビットバッファ202の占有量 B_m の変化を示す。ビットバッファ202の占有量 B_m はビットレート RB をグラフの傾きとして上昇する。ビットレート RB は、シーケンスの先頭に付くシーケンスヘッダの BR (Bit Rate) に従って式(1)に示すように規定される。また、AVバーサ11から転送されてくるビデオストリームのピクチャレート RP はシーケンスヘッダの PR (Picture Rate) によって規定される。そして、ビットバッファ202の容量 B は、シーケンスヘッダの VBV (Vbv[Video Bufferring Verifier] Buffer Size) に従って式(2)に示すように規定される。そして、1フレーム期間毎に、デコードコア回路204がそのときデコードしようとする1ピクチャ分のビデオストリームが、ビットバッファ202から一気に読み出される。ここで、1フレーム期間にビットバッファ202に入力されるビデオストリームのデータ量 X は、ビットレート RB およびピクチャレート RP に従って式(3)に示すように規定される。従って、ビットバッファ202から1ピクチャ分のビデオストリームが一気に読み出された直後のビットバッファ202の占有量 B_m ($=B_0 \sim B_6$) は、データ量 X とビットバッファ202の容量 B とに基づいて、式(4)に示す条件を満たすように規定される。

【0118】

$$RB = 400 \times BR \quad \dots\dots (1)$$

$$B = 16 \times 1024 \times VBV \quad \dots\dots (2)$$

$$X = RB / RP \quad \dots\dots (3)$$

$$0 < B_m < B - X = B - (RB / RP) \quad \dots\dots (4)$$

式(4)に示す条件を満たすようにビットバッファ202の占有量 B_m が規定されていれば、ビットバッファ202がオーバーフローしたりアンダーフローしたりすることはない。逆に言えば、ビットバッファ202の占有量 B_m が閾値 $(B - X)$ を越えると、次の1フレーム期間にビットバッファ202に入力されるビデオストリー *

22

*ムによってビットバッファ202がオーバーフローする可能性が極めて高くなる。

【0119】ビデオデコーダ12では、通常の再生時において、式(4)が満たされるように、ビットレート RB 、ピクチャレート RP 、容量 B の各値が規定されている。つまり、式(2)に示すようにビットバッファ202の容量 B を設定しておけば、ピクチャスキップ回路206の接続をノード206a側に固定しておいたとしても、理想的な状態ではビットバッファ202がオーバーフローしたりアンダーフローしたりすることはない。

【0120】従って、通常の再生時において、ビットバッファ202から1ピクチャ分のデータが一気に読み出された直後の占有量 B_m ($=B_0 \sim B_4$) は、閾値 B_{th1} に基づいて、式(5)に示す条件を満たすように規定される。尚、閾値 B_{th1} は、式(4)に基づいて、式(6)に示すように設定される。

【0121】

$$0 < B_m < B_{th1} < B \quad \dots\dots (5)$$

$$B_{th1} = B - X = B - (RB / RP) \quad \dots\dots (6)$$

ところで、実際の状態では、式(2)に示すようにビットバッファ202の容量 B を設定しておいても、ピクチャスキップ回路206の接続をノード206a側に固定しておく、ビットバッファ202がオーバーフローする恐れがある。

【0122】しかし、ビデオデコーダ12では、通常の再生時において、ビットバッファ202の占有量 B_m が閾値 B_{th1} を越えた場合、ビットバッファ202がオーバーフローする恐れがあると判定される。すると、ビットバッファ202の占有量 B_m が閾値 B_{th1} を下回るまで、ビットバッファ202から適宜なピクチャ分のビデオストリームが読み出される。そして、ピクチャスキップ回路206はノード206b側に接続され、そのビットバッファ202から読み出された適宜なピクチャ分のビデオストリームは全てスキップされる。従って、ビデオデコーダ12によれば、通常の再生時において、ビットバッファ202がオーバーフローすることはない。

【0123】高速再生時におけるビットバッファ202の占有量 B_m はビットレート $n \times RB$ をグラフの傾きとして上昇する。例えば、2倍速再生時におけるビットバッファ202の占有量 B_m はビットレート $2 \times RB$ をグラフの傾きとして上昇する。

【0124】従って、高速再生時において、ビットバッファ202から1ピクチャ分のデータが一気に読み出された直後の占有量 B_m ($=B_0 \sim B_4$) は、閾値 B_{thn} に基づいて、式(7)に示す条件を満たすように規定される。尚、閾値 B_{thn} は式(8)に示すように設定される。

【0125】

$$0 < B_m < B_{thn} \quad \dots\dots (7)$$

$$B_{thn} = B - n \times X = B - (n \times RB / RP) \quad \dots\dots (8)$$

高速再生時においては、ビットバッファ202の占有量 B_m が閾値 B_{thn} を越えた場合、ビットバッファ202がオーバーフローする恐れがあると判定される。例えば、2倍速再生時には占有量 B_m が閾値 B_{th2} ($=B - (2 \times RB / RP)$) を越えた場合、3倍速再生時には占有量 B_m が閾値 B_{th3} ($=B - (3 \times RB / RP)$) を越えた場合に、ビットバッファ202がオーバーフローする恐れがあると判定される。すると、ビットバッファ202の占有量 B_m が閾値 B_{thn} を下回るまでビットバッファ202から適宜なピクチャ分のビデオストリームが読み出され、そのビデオストリームは全てスキップされる。従って、ビデオデコーダ12によれば、高速再生時において、ビットバッファ202がオーバーフローすることはない。

【0126】デコードコア回路204において任意のピクチャをデコードしている途中でビットバッファ202がオーバーフローすると、デコード処理中のピクチャのビットバッファ202に残っている部分に対して、新たに入力されたビデオストリームが上書きされる。その結果、デコード処理中のピクチャのビットバッファ202に残っている部分が破壊されて失われる。すると、デコードコア回路204では、そのピクチャのデコードを完了することが不可能になり、そのピクチャのビデオ信号を生成することができなくなる。従って、デコードコア回路204において任意のピクチャをデコードしている途中でビットバッファ202がオーバーフローすることは絶対に避けなければならない。

【0127】そのため、ビットバッファ202がオーバーフローする恐れがあるかどうかの判定は、デコードコア回路204において任意のピクチャのデコードを開始する前に行う必要がある。より正確には、ピクチャヘッダ検出回路203がピクチャヘッダを検出した時点で、ビットバッファ202がオーバーフローする恐れがあるかどうかを判定し、そのピクチャをピクチャスキップ回路206を介してスキップするかどうかを決定する必要がある。

【0128】ところで、1つのピクチャのデータ量は0~40バイトであるが、そのデータ量はデコードコア回路204においてデコードが終了した時点でないとわからない。また、1つのピクチャのデコード処理時間は、そのピクチャのデータ量やデコードコア回路204の動作速度によって異なるが、通常、1フレーム期間の1/3~3/4程度である。

【0129】ビットバッファ202から読み出されたピクチャのデータ量が0バイトの場合、そのピクチャの読み出し前後でビットバッファ202の占有量 B_m は変化しないため、そのピクチャをスキップしたとしてもオーバーフローを回避することはできない。逆に言えば、ビットバッファ202から読み出されたピクチャのデータ量が0バイトの場合でも、ビットバッファ202に十分

な空き容量があればオーバーフローすることはない。

【0130】そこで、1フレーム期間にビットバッファ202に入力されるビデオストリームのデータ量分の空き容量を、ビットバッファ202に確保しておく。そうすれば、ビットバッファ202から読み出されたピクチャのデータ量が0バイトの場合でもオーバーフローすることはない。

【0131】1フレーム期間にビットバッファ202に入力されるビデオストリームのデータ量は、($n \times X = n \times RB / RP$)になる。ビットバッファ202の空き容量がこのデータ量以上であればオーバーフローすることはない。従って、式(8)に示すように閾値 B_{thn} を設定しておけば、ビットバッファ202のオーバーフローを確実に回避することができる。

【0132】すなわち、判定回路205は、ピクチャヘッダ検出回路203がピクチャヘッダを検出した時点でビットバッファ202の空き容量をチェックし、十分な空き容量 ($n \times X = n \times RB / RP$) が確保されているかどうかを判定する。十分な空き容量が確保されていないければ、そのピクチャヘッダに基づいて制御コア回路207がビットバッファ202から読み出したピクチャを、ピクチャスキップ回路206を介してスキップする。続いて、判定回路205は、ピクチャヘッダ検出回路203が次のピクチャヘッダを検出した時点で、再びビットバッファ202の空き容量をチェックする。これらの処理に要する時間は、デコードコア回路204のデコード処理時間に比べてはるかに短いため、ビットバッファ202に十分な空き容量が確保できてからデコードコア回路204のデコード処理を開始しても十分に間に合う。

【0133】ところで、ピクチャヘッダ検出回路203がピクチャヘッダを検出した時点や、デコードコア回路204がデコードを開始した後に、ビットバッファ202がアンダーフローすることがある。この場合は、ビデオストリームがビットバッファ202に入力され次第、ビットバッファ202から1ピクチャ分のビデオストリームを逐次読み出せばよいため、特に問題とはならない。

【0134】以上詳述したように、ビデオデコーダ12によれば、以下に示す効果を得ることができる。

①通常の再生時において、ビットバッファ202のオーバーフローを回避することができる。

【0135】②高速再生時において、ビットバッファ202のオーバーフローを回避することができる。

③判定回路205およびピクチャスキップ回路206を設けることにより、ビットバッファ202のオーバーフローを回避することができる。上記したように判定回路205およびピクチャスキップ回路206の制御は簡単であるため、制御コア回路207はマイクロコンピュータを用いて構成する必要がない。そして、各回路203

～207を1チップのLSIに搭載した場合には、ビデオデコーダ12を小型化することができる。

【0136】④ピクチャスキップ回路206のノード206b側からスキップされるビデオストリームは、ピクチャ単位となる。そのため、デコードコア回路204へ転送されるピクチャの途中でデータが途切れることはない。従って、デコードコア回路204では、IピクチャだけでなくPピクチャやBピクチャについてもデコード可能になる。その結果、ディスプレイ22で再生される動画に生じるコマ落ちが少なくなる。そのため、2～4倍という比較的遅い高速再生時において、数コマ/秒の表示が可能になる。従って、高速再生時における動画の動きを滑らかにして画質を大幅に向上させることができる。

【0137】ところで、上記したビデオデコーダ12において、式(9)に示す規定を満たすように、2つの閾値B2thn、B3thnを設定してもよい。尚、各閾値B2thn、B3thnの値は、上記のように再生速度に応じて設定されると共に、ディスプレイ22で再生される動画の画質を実際に検討して適宜に設定すればよい。

【0138】 $0 < B3thn < B2thn < B$ …… (9)
判定回路205は、ビットバッファ202の占有量Bmと各閾値Bthn、B2thnとを比較し、占有量Bmが式(10)～(12)に示すどの領域に含まれるかを判定する。

【0139】 $Bm < B3thn$ …… (10)

$B3thn < Bm < B2thn$ …… (1.1)

$B2thn < Bm$ …… (1.2)

判定回路205は、式(10)に示すように、ビットバッファ202の占有量Bmが閾値B3thnを越えない場合には、ビットバッファ202がオーバーフローする恐れがなく正常であると判定する。この場合、制御コア回路207は、ビットバッファ202から1ピクチャ分のビデオストリームを読み出す。そして、制御コア回路207は、ピクチャスキップ回路206をノード206a側に接続し、そのビットバッファ202から読み出されたピクチャをデコードコア回路204へ転送させる。

【0140】判定回路205は、式(1.2)に示すように、ビットバッファ202の占有量Bmが閾値B2thnを越え且つ閾値Bthnを越えない場合に、ビットバッファ202から読み出されたピクチャがIピクチャまたはPピクチャならば、第1のフラグを立てる。また、式(1.1)に示すように、ビットバッファ202の占有量Bmが閾値B3thnを越え且つ閾値B2thnを越えない場合に、ビットバッファ202から読み出されたピクチャがPピクチャならば、第2のフラグを立てる。第1または第2のフラグが立っている場合、式(10)に示す場合でも、制御コア回路207は、ビットバッファ202から読み出されたピクチャがBピクチャならば、ピクチャスキップ回路206をノード206b側に接続し、そのピ

クチャをスキップさせる。

【0141】図13に、2つの閾値B2thn、B3thnを設定した場合におけるビットバッファ202の占有量Bmの変化を示す。占有量Bmが閾値B3thnを越えた場合、ビットバッファ202から読み出されたピクチャがBピクチャであればデコードせずにスキップする(図示※1)。ここで、Bピクチャのスキップ後に占有量Bmがまだ閾値B3thnを越えていても、ビットバッファ202から次に読み出されたピクチャがIピクチャまたはPピクチャであればデコードする(図示※2)。

【0142】占有量Bmが閾値B3thnを越えた場合でも、ビットバッファ202から読み出されたピクチャがIピクチャまたはPピクチャであればデコードする(図示※3)。ここで、IピクチャまたはPピクチャのデコード後に占有量Bmがまだ閾値B3thnを越えている場合、ビットバッファ202から次に読み出されたピクチャがBピクチャであればデコードせずにスキップする(図示※4)。このBピクチャのスキップは、占有量Bmが閾値B3thnを下回るまで繰り返し行う(図示※5)。

【0143】占有量Bmが閾値B2thnを越えた場合、ビットバッファ202から読み出されたピクチャがIピクチャまたはPピクチャであれば、判定回路205は第1のフラグを立てる(図示※6)。第1のフラグが立っている場合、ビットバッファ202から次に読み出されたピクチャがBピクチャであれば、占有量Bmが閾値B3thnを下回っていても、そのBピクチャをスキップする(図示※7)。

【0144】占有量Bmが閾値B3thnを越え且つ閾値B2thnを越えない場合、ビットバッファ202から読み出されたピクチャがPピクチャであれば、判定回路205は第2のフラグを立てる(図示※8)。第2のフラグが立っている場合、ビットバッファ202から次に読み出されたピクチャがBピクチャであれば、占有量Bmが閾値B3thnを下回っていても、そのBピクチャをスキップする(図示※9)。

【0145】占有量Bmが閾値B3thnを越え且つ閾値B2thnを越えない場合、ビットバッファ202から読み出されたピクチャがIピクチャのときには、判定回路205は第2のフラグを立てない(図示※10)。第2のフラグが立っていない場合、占有量Bmが閾値B3thnを下回っていれば、ビットバッファ202から次に読み出されたピクチャがBピクチャであってもデコードする。

【0146】以上のように、2つの閾値B2thn、B3thnを設定した場合には、上記したビデオデコーダ12の効果①～③に加えて、以下の効果を得ることができる。

④ビットバッファ202の占有量Bmが閾値B3thnを越え且つ閾値Bthnを越えない場合、IピクチャおよびPピクチャを可能な限りデコードすると共に、Bピクチャを優先してスキップする。

れたMPEGシステムストリームを、MPEGオーディオストリームとMPEGビデオストリームとに分離するデマルチプレクサ(13)を備えたMPEG再生装置。

【0162】このようにすれば、オーディオデコーダへオーディオストリームを、ビデオデコーダへビデオストリームをそれぞれ転送することができる。

【0163】

【発明の効果】請求項1、3～6のいずれか1項に記載の発明によれば、可変速再生時においても自然で聞き易い音声を再生することが可能なMPEGオーディオ再生装置を提供することができる。

【0164】請求項2または請求項7に記載の発明によれば、可変速再生時においても自然で聞き易い音声を再生することが可能なMPEGオーディオ再生装置とMPEGビデオデコーダとを備えたMPEG再生装置を提供することができる。

【0165】請求項8～11のいずれか1項に記載の発明によれば、可変速再生時においても自然で聞き易い音声を再生することが可能なMPEGオーディオ再生装置とMPEGビデオデコーダとを備え、音声と動画との時間ずれを低減することが可能なMPEG再生装置を提供することができる。

【図面の簡単な説明】

【図1】第1実施形態のブロック回路図。

【図2】第1実施形態の要部ブロック回路図。

【図3】第1実施形態の作用を説明するための模式図。

【図4】第1実施形態の作用を説明するための模式図。

【図5】第2実施形態の要部ブロック回路図。

【図6】第3実施形態の要部ブロック回路図。

【図7】第4実施形態の要部ブロック回路図。

【図8】第5実施形態の要部ブロック回路図。

【図9】第6実施形態の要部ブロック回路図。

【図10】第7実施形態の要部ブロック回路図。

【図11】第8実施形態の要部ブロック回路図。

【図12】MPEGビデオデコーダの要部ブロック回路図。

【図13】MPEGビデオデコーダの動作を説明するためのグラフ。

【図14】MPEGビデオデコーダの動作を説明するためのグラフ。

【符号の説明】

1…MPEGオーディオ再生装置

2…話速変換手段としての再生速度検出回路

3…MPEGオーディオデコーダ

4…話速変換手段としての話速変換処理回路

12…MPEGビデオデコーダ

21…記録媒体

32…リングメモリ

33…検出手段としてのアップダウンカウンタ

41…音声判別部

42…無音削除挿入部

43…時間軸圧縮伸長部

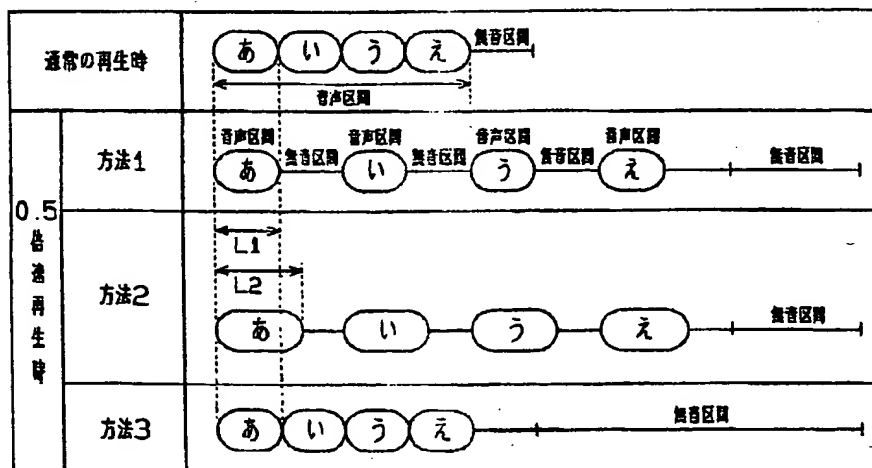
51…インデックス付加回路

52…インデックス検出回路

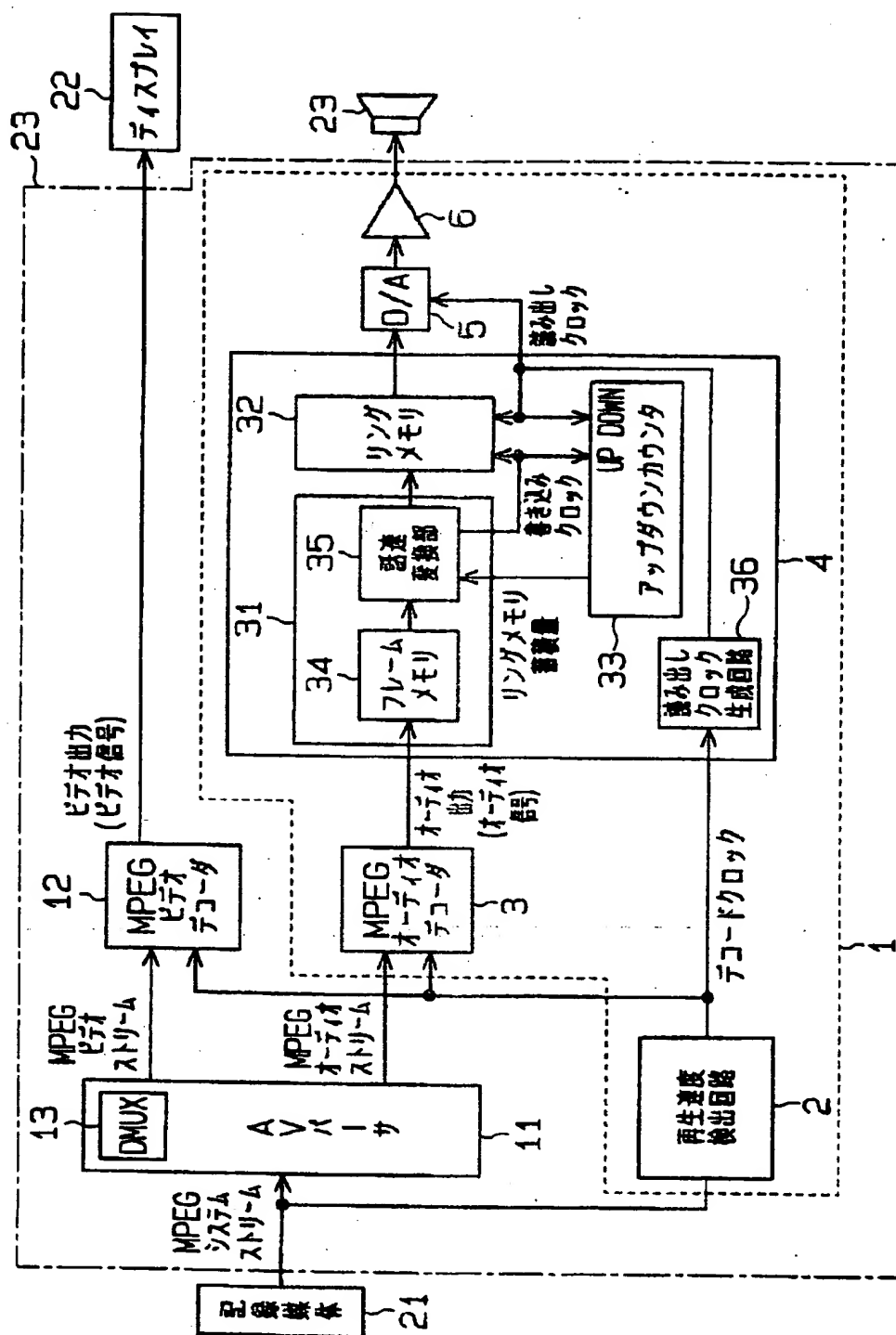
53、55…遅延時間検出回路

54…制御回路

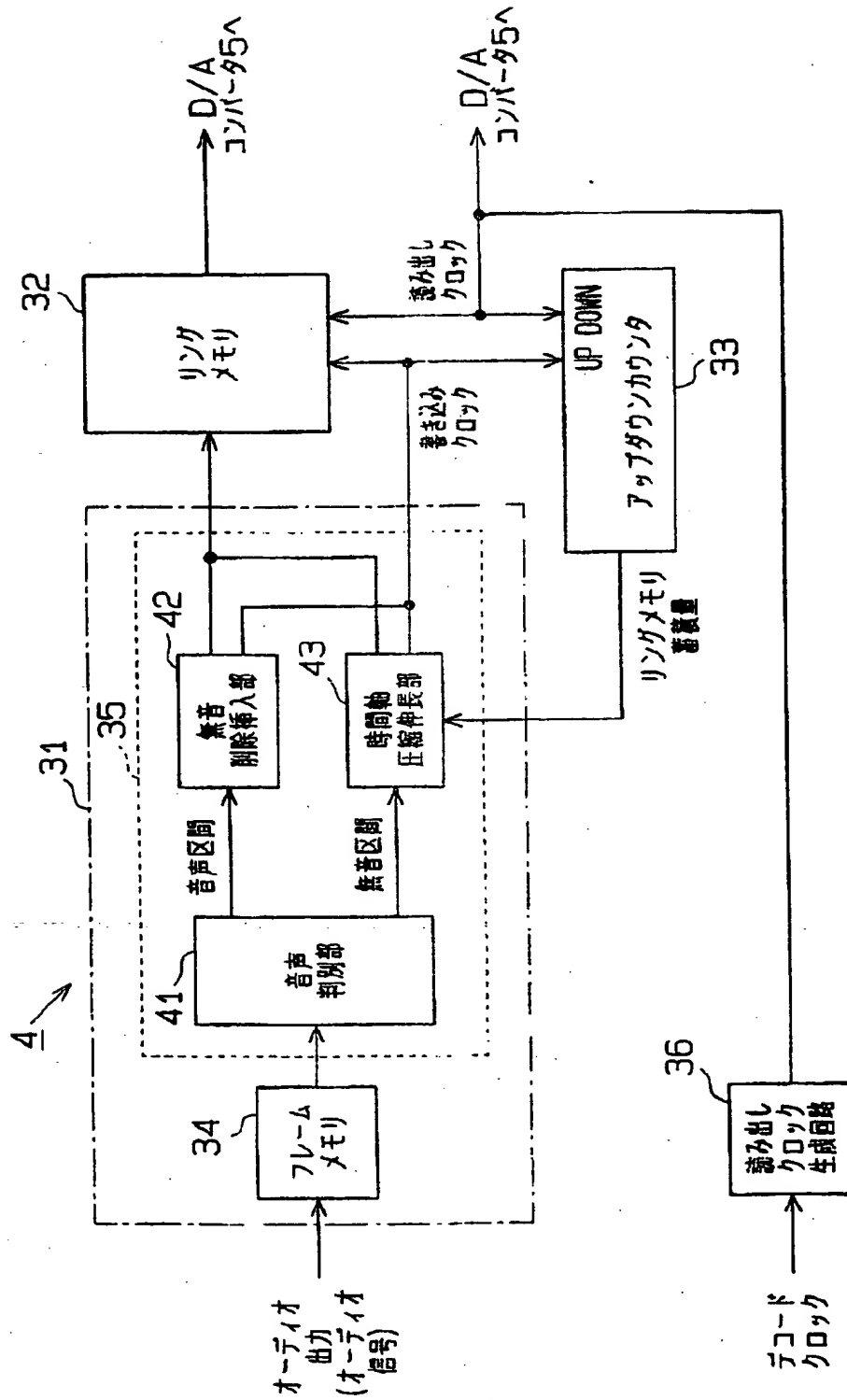
【図3】



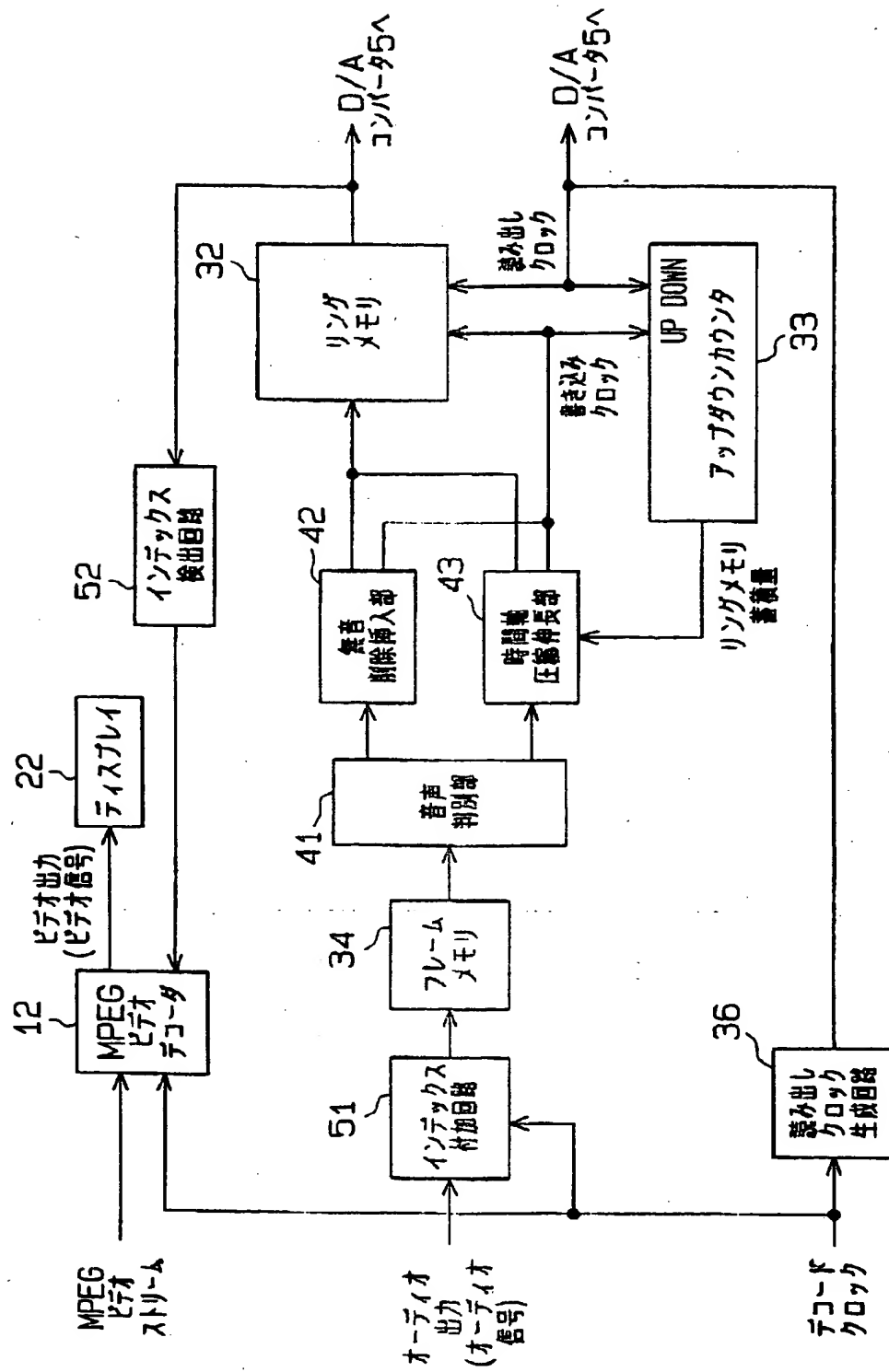
【図1】



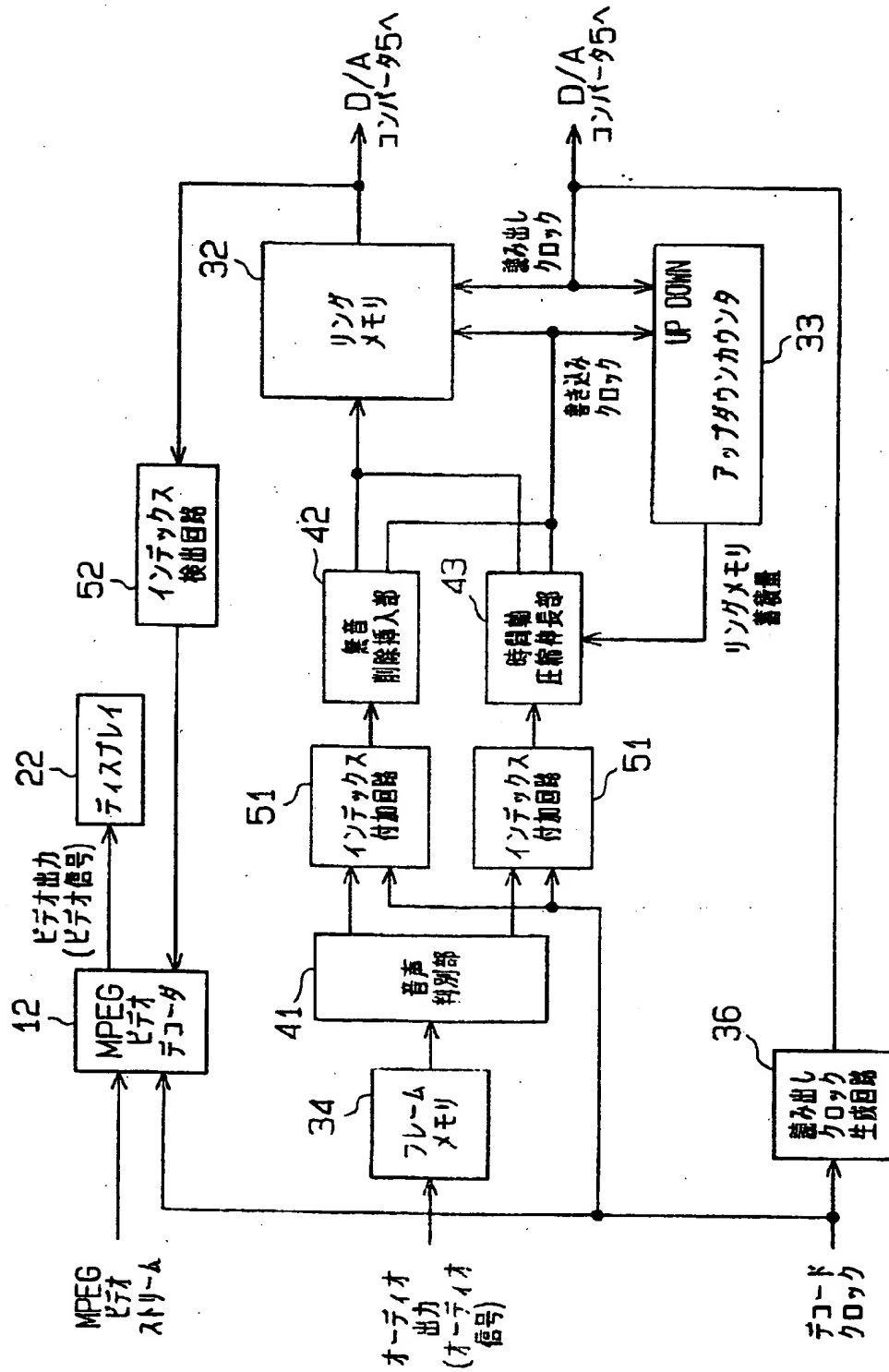
【図2】



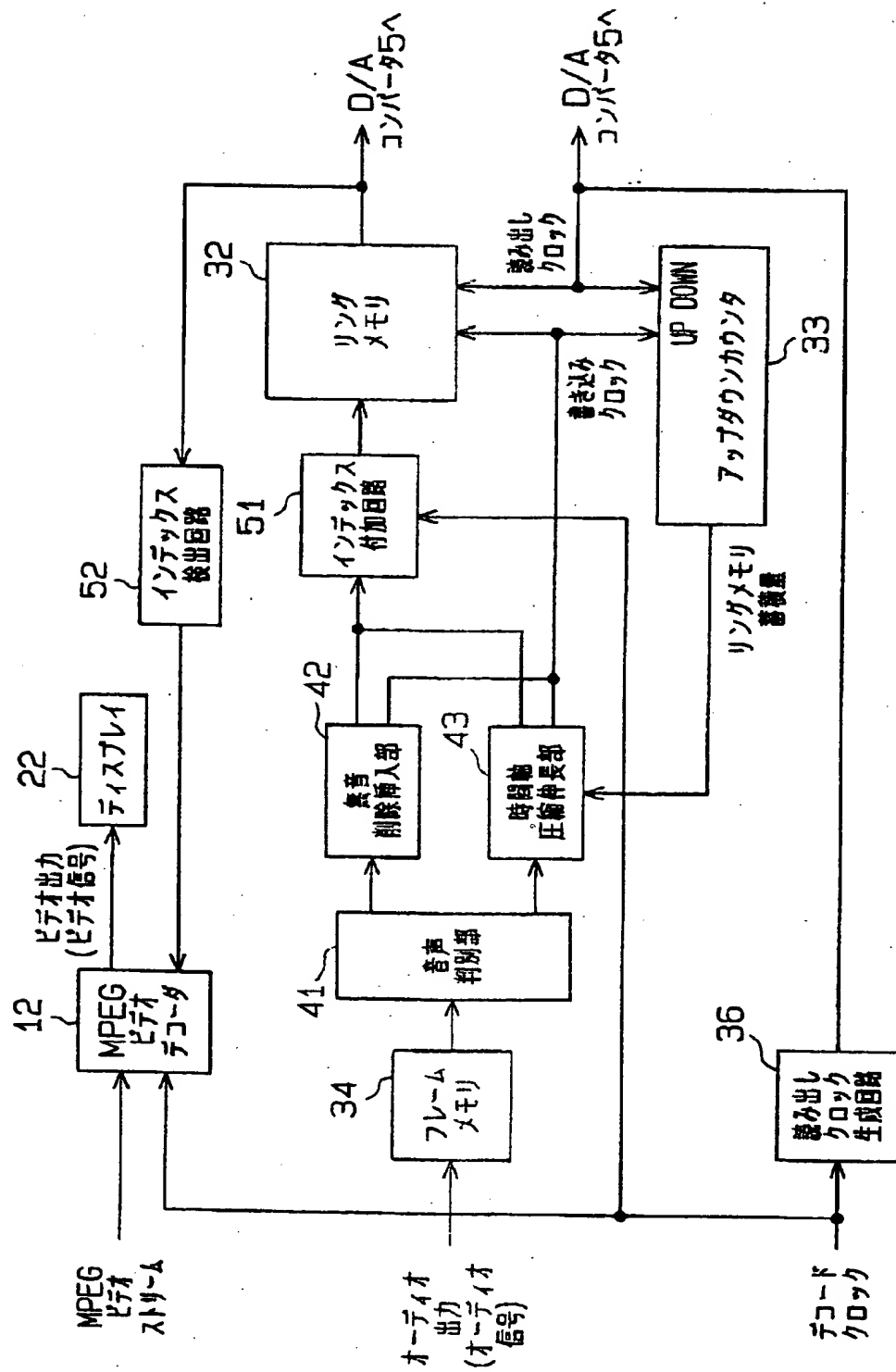
【図5】



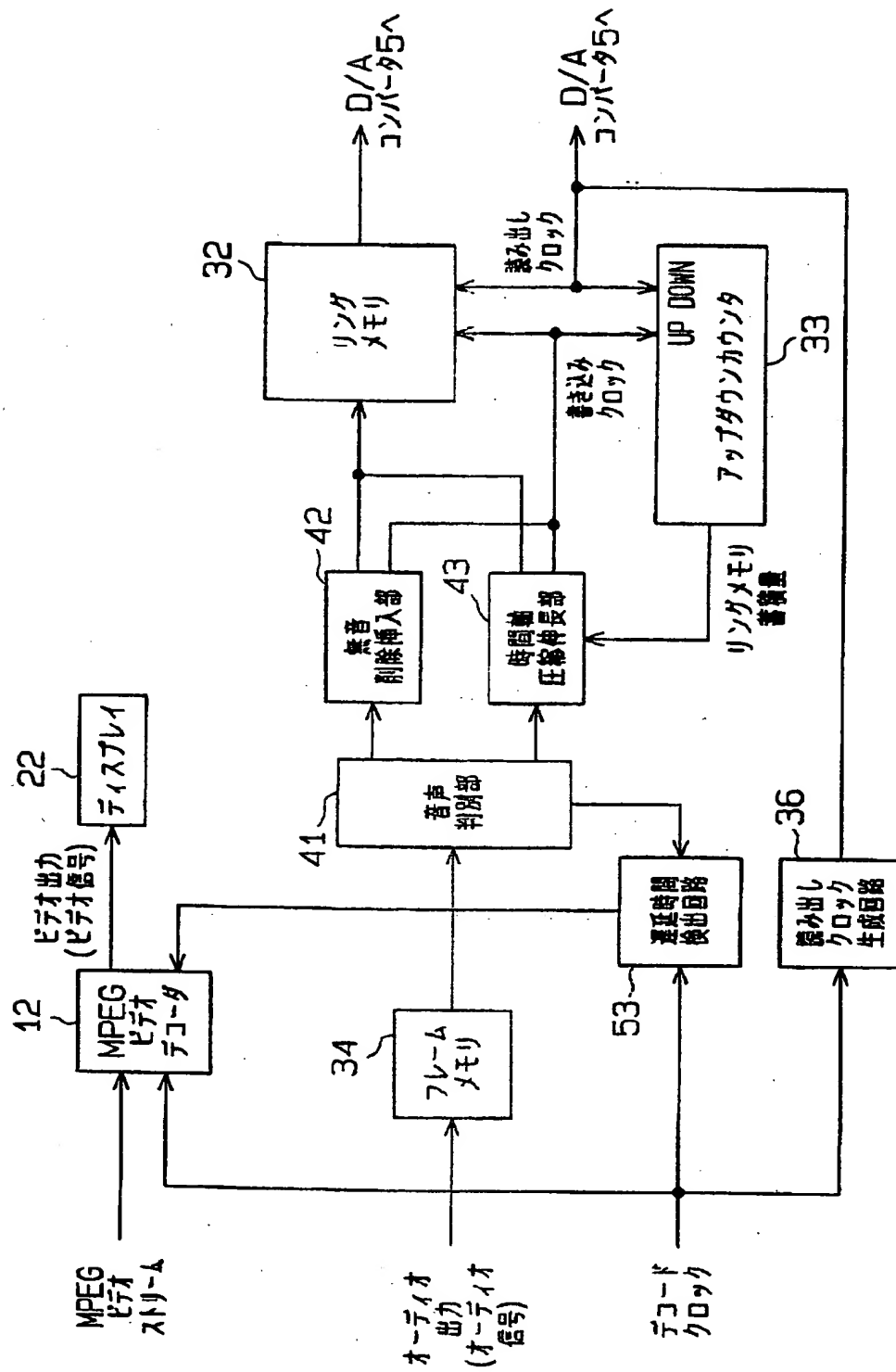
【図7】



【図8】



【図9】



【図11】

